

ENHANCED CHOICE OF THE PARAMETERS IN AN ITERATIVELY REGULARIZED NEWTON- LANDWEBER ITERATION IN BANACH SPACE

BARBARA KALTENBACHER

Alpen-Adria-Universität Klagenfurt
Universitätstrasse 65–67
9020 Klagenfurt, Austria

IVAN TOMBA

Università di Bologna
Piazza Porta S. Donato, 5
40127 - Bologna, Italy

(Communicated by the associate editor name)

ABSTRACT. This paper is a close follow-up of [9] and [10], where Newton-Landweber iterations have been shown to converge either (unconditionally) without rates or (under an additional regularity assumption) with rates. The choice of the parameters in the method were different in each of these two cases. We now found a unified and more general strategy for choosing these parameters that enables both convergence and convergence rates. Moreover, as opposed to the previous one, this choice yields strong convergence as the noise level tends to zero, also in the case of no additional regularity. Additionally, the resulting method appears to be more efficient than the one from [9], as our numerical tests show.

1. Introduction. Regularization of inverse problems in Banach spaces is a field of highly active research, cf., e.g., [2, 16, 8] for variational regularization and, e.g., [1, 6, 10, 11, 13, 15, 21], for iterative methods. The main reason for this lies in the fact that extension of the scope from Hilbert to general Banach spaces better allows to formulate requirements on the searched for solution and to describe realistic noise models.

We will here especially concentrate on a combination of a Newton-type strategy with Landweber iterations to approximate the Newton step, which leads to a fully explicit iteration, cf. [9, 10].

To formulate the method, consider a nonlinear ill-posed operator equation

$$F(x) = y \tag{1}$$

where F maps between Banach spaces X and Y . The given data y^δ are typically contaminated by noise, and we are going to assume that the noise level δ in

$$\|y - y^\delta\| \leq \delta \tag{2}$$

2010 *Mathematics Subject Classification.* Primary 65J20, Secondary 65M32.

Key words and phrases. regularization, nonlinear inverse problems, Banach space, Newton's method, Landweber iteration.

is known. In the following, x_0 is some initial guess and we will assume that a solution x^\dagger to (1) exists. For some $p, r \in (1, \infty)$, we will make use of the duality mappings $J_p^X(x) := \partial \left\{ \frac{1}{p} \|x\|^p \right\}$ from X to its dual X^* , and $J_r^Y(y) := \partial \left\{ \frac{1}{r} \|y\|^r \right\}$ from Y to Y^* , respectively. While under the assumptions we will make on X , the mapping J_p^X will in fact be single valued, this will not necessarily be the case for J_r^Y and we will denote by j_r^Y a single valued selection from J_r^Y . Therewith, we consider a combination of the iteratively regularized Gauss-Newton method with an iteratively regularized Landweber method for approximating the Newton step, using some initial guess x_0 and starting from some x_0^δ (that need not necessarily coincide with x_0)

$$\begin{aligned}
& \text{For } n = 0, 1, 2 \dots \text{ do} \\
& \quad u_{n,0} = 0 \\
& \quad z_{n,0} = x_n^\delta \\
& \quad \text{For } k = 0, 1, 2 \dots, k_n - 1 \text{ do} \\
& \quad \quad u_{n,k+1} = u_{n,k} - \alpha_{n,k} J_p^X(z_{n,k} - x_0) \\
& \quad \quad \quad - \omega_{n,k} F'(x_n^\delta)^* j_r^Y(F'(x_n^\delta)(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta) \\
& \quad \quad z_{n,k+1} = x_0 + J_p^{X^{-1}} \left(J_p^X(x_n^\delta - x_0) + u_{n,k+1} \right) \\
& \quad x_{n+1}^\delta = z_{n,k_n}.
\end{aligned} \tag{3}$$

It is clear that the choice of the parameters $\alpha_{n,k}$, $\omega_{n,k}$, k_n , and of the overall stopping index $n = n_*$ crucially influences the stability and efficiency of the method. While in [9], [10] only either convergence or convergence rates have been established with disjoint parameter choices for each of these two cases, the aim of this paper is to provide a unified parameter choice strategy for this method that allows to show both unconditional convergence and convergence rates under additional regularity assumptions on the solution. Moreover, by using an appropriate choice of the stopping index for the inner iteration, differently from [9], [10] we can show continuous dependence of the iterates x_n on the data y^δ for each fixed outer iteration index n and therewith are able to prove strong convergence. Finally, the new parameter choice appears to enhance efficiency as compared to [9], as the numerical tests below show.

The remainder of this paper is organized as follows. In Section 2 we provide some preliminaries. The parameter choice as well as convergence results are derived and formulated in Section 3. Section 4 shows some numerical tests for a coefficient identification problem in an elliptic PDE in one and two space dimensions.

2. Preliminaries and assumptions. Throughout this paper we will assume that X is smooth, which means that the duality mapping is single-valued, and moreover, that X is s -convex for some $s \in [p, \infty)$, which implies

$$D_p(x, y) \geq \frac{c_{p,s}}{(\|x\| + \|y\|)^{s-p}} \|x - y\|^s \tag{4}$$

for some constant $c_{p,s} > 0$, cf. Corollary 2.61 in [21]. Here, $D_p(x, y)$ denotes the Bregman distance

$$D_p(\tilde{x}, x) = \frac{1}{p} \|\tilde{x}\|^p - \frac{1}{p} \|x\|^p - \langle j_p^X(x), \tilde{x} - x \rangle_{X^*, X}$$

(where $j_p^X(x)$ denotes a single valued selection of $J_p^X(x)$). We will also make use of its shifted version

$$D_p^{x_0}(\tilde{x}, x) := D_p(\tilde{x} - x_0, x - x_0).$$

As a consequence of the above assumptions, X is reflexive and we also have

$$D_{p^*}(x^*, y^*) \leq C_{p^*, s^*} \|x^* - y^*\|^{s^*} ((pD_{p^*}(J_{p^*}^{X^*}(x^*), 0))^{1-\frac{s^*}{p^*}} + \|x^* - y^*\|^{p^*-s^*}), \quad (5)$$

for some C_{p^*, s^*} , where s^* denotes the dual index $s^* = \frac{s}{s-1}$, cf. (4) in [9]. Under these assumptions, the duality mapping is bijective and $J_p^{-1} = J_{p^*}^{X^*}$, the latter denoting the (by s -convexity also single-valued) duality mapping on the dual X^* of X . We will also make use of the identities

$$D_p(x, y) = D_p(x, z) + D_p(z, y) + \langle J_p^X(z) - J_p^X(y), x - z \rangle_{X^*, X} \quad (6)$$

and

$$D_p(y, x) = D_{p^*}(J_p^X(x), J_p^X(y)). \quad (7)$$

For more details on the geometry of Banach spaces we refer, e.g., to [20], [21] and the references therein.

The assumptions on the forward operator besides a condition on the domain

$$\mathcal{B}_\rho^D(x^\dagger) \subseteq \mathcal{D}(F) \quad (8)$$

include a structural condition on its degree of nonlinearity. For simplicity of exposition we restrict ourselves to the tangential cone condition

$$\|F(\tilde{x}) - F(x) - F'(x)(\tilde{x} - x)\| \leq \eta \|F(\tilde{x}) - F(x)\|, \quad \tilde{x}, x \in \mathcal{B}_\rho^D(x^\dagger), \quad (9)$$

and mention in passing that this could be extended to a more general condition on the degree of nonlinearity (cf. [5]) as in [12]. Here F' is not necessarily the Fréchet derivative of F but just a linearization of F satisfying the Taylor remainder estimate (9). Additionally, we assume that F' and F are uniformly bounded on $\mathcal{B}_\rho^D(x^\dagger)$.

Here

$$\mathcal{B}_\rho^D(x^\dagger) = \{x \in X \mid D_p^{x_0}(x^\dagger, x) \leq \rho^2\}$$

with $\rho > 0$ such that $x_0 \in \mathcal{B}_\rho^D(x^\dagger)$. By distinction between the cases $\|x - x_0\| < 2\|x^\dagger - x_0\|$ and $\|x - x_0\| \geq 2\|x^\dagger - x_0\|$ and the second triangle inequality we obtain from (4) that

$$\mathcal{B}_\rho^D(x^\dagger) \subseteq \mathcal{B}_{\bar{\rho}}^{\|\cdot\|}(x_0) = \{x \in X \mid \|x - x_0\| \leq \bar{\rho}\} \quad (10)$$

with $\bar{\rho} = \max\{2\|x^\dagger - x_0\|, \left(\frac{2^p 3^{s-p} \rho^2}{c_{p,s}}\right)^{1/p}\}$

For obtaining convergence rates we impose a variational inequality (or variational source condition)

$$\exists \beta > 0 : \forall x \in \mathcal{B}_\rho^D(x^\dagger)$$

$$|\langle J_p^X(x^\dagger - x_0), x - x^\dagger \rangle_{X^* \times X}| \leq \beta D_p^{x_0}(x^\dagger, x)^{\frac{1-2\nu}{s}} \|F(x) - F(x^\dagger)\|^{2\nu \frac{\lambda(s,\nu)}{\lambda(2,\nu)}} \quad (11)$$

with s the parameter of smoothness of X ,

$$\lambda(s, \nu) = 1 - \frac{1-2\nu}{s} \quad (12)$$

and $\nu \in (0, \frac{1}{2}]$. Condition (11) corresponds to a source condition $x^\dagger - x_0 \in \mathcal{R}((F'(x^\dagger)^* F'(x^\dagger))^\nu)$ in the special case of Hilbert spaces (where $s = 2$), cf., e.g., [5]. Note that (11) is stronger for larger ν and (11) always holds for $\nu = 0$ with

$$\beta = \beta_0 = \|x^\dagger - x_0\|^{p-1} \left(\frac{(\bar{\rho} + \|x^\dagger - x_0\|)^{s-p}}{c_{p,s}} \right)^{\frac{1}{s}}, \quad (13)$$

due to (4), (10). The case $\nu = 0$ can be identified with the situation that no additional regularity (i.e., (11) with $\nu > 0$) is known to hold.

3. Error estimates and parameter choice. We here first of all follow the lines of Section 2 in [9]. Some of the estimates are the same as in [9] (and will be repeated here only for convenience of the reader). Some of them are different, though and therewith enable different parameter choice strategies.

For any $n \in \mathbb{N}$ we have

$$\begin{aligned}
& D_p^{x_0}(x^\dagger, z_{n,k+1}) - D_p^{x_0}(x^\dagger, z_{n,k}) \\
&= D_p^{x_0}(z_{n,k}, z_{n,k+1}) + \underbrace{\langle J_p^X(z_{n,k+1} - x_0) - J_p^X(z_{n,k} - x_0), z_{n,k} - x^\dagger \rangle_{X^* \times X}}_{=u_{n,k+1}-u_{n,k}} \\
&= \underbrace{D_p^{x_0}(z_{n,k}, z_{n,k+1})}_{(I)} \\
&\quad - \underbrace{\omega_{n,k} \langle j_r^Y(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta), A_n(z_{n,k} - x^\dagger) \rangle_{Y^* \times Y}}_{(II)} \\
&\quad - \underbrace{\alpha_{n,k} \langle J_p^X(x^\dagger - x_0), z_{n,k} - x^\dagger \rangle_{X^* \times X}}_{(III)} \\
&\quad - \underbrace{\alpha_{n,k} \langle J_p^X(z_{n,k} - x_0) - J_p^X(x^\dagger - x_0), z_{n,k} - x^\dagger \rangle_{X^* \times X}}_{(IV)}, \tag{14}
\end{aligned}$$

where we abbreviate

$$A_n = F'(x_n^\delta), \quad b_n = y^\delta - F(x_n^\delta). \tag{15}$$

Assuming that $z_{n,k} \in \mathcal{B}_\rho^D(x^\dagger)$, we now estimate each of the terms on the right hand side separately.

By (5) and (7) we have for the term (I)

$$\begin{aligned}
& D_p^{x_0}(z_{n,k}, z_{n,k+1}) \\
&\leq C_{p^*,s^*} \| \underbrace{J_p^X(z_{n,k+1} - x_0) - J_p^X(z_{n,k} - x_0)}_{=u_{n,k+1}-u_{n,k}} \|^{s^*} \\
&\quad \cdot \left((p\rho^2)^{1-\frac{s^*}{p^*}} + \| J_p^X(z_{n,k+1} - x_0) - J_p^X(z_{n,k} - x_0) \|^{p^*-s^*} \right) \\
&= C_{p^*,s^*} (p\rho^2)^{1-\frac{s^*}{p^*}} \| \alpha_{n,k} J_p^X(z_{n,k} - x_0) + \omega_{n,k} A_n^* j_r^Y(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta) \|^{s^*} \\
&\quad + C_{p^*,s^*} \| \alpha_{n,k} J_p^X(z_{n,k} - x_0) + \omega_{n,k} A_n^* j_r^Y(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta) \|^{p^*} \\
&\leq 2^{s^*-1} C_{p^*,s^*} (p\rho^2)^{1-\frac{s^*}{p^*}} \alpha_{n,k}^{s^*} \| z_{n,k} - x_0 \|^{(p-1)s^*} \\
&\quad + 2^{s^*-1} C_{p^*,s^*} (p\rho^2)^{1-\frac{s^*}{p^*}} \omega_{n,k}^{s^*} \| A_n^* j_r^Y(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta) \|^{s^*} \\
&\quad + 2^{p^*-1} C_{p^*,s^*} \alpha_{n,k}^{p^*} \| z_{n,k} - x_0 \|^{(p-1)p^*} \\
&\quad + 2^{p^*-1} C_{p^*,s^*} \omega_{n,k}^{p^*} \| A_n^* j_r^Y(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta) \|^{p^*} \\
&\leq C_{p^*,s^*} \left((p\rho^2)^{1-\frac{s^*}{p^*}} \bar{\rho}^{(p-1)s^*} 2^{s^*-1} \alpha_{n,k}^{s^*} + \bar{\rho}^{p^*} 2^{p^*-1} \alpha_{n,k}^{p^*} \right) + \varphi(\omega_{n,k} \tilde{t}_{n,k}) \tag{16}
\end{aligned}$$

where we have used the triangle inequality in X^* and X , the inequality

$$(a+b)^\lambda \leq \begin{cases} 2^{\lambda-1}(a^\lambda + b^\lambda) & \text{if } \lambda \geq 1 \\ (a^\lambda + b^\lambda) & \text{if } 0 \leq \lambda \leq 1 \end{cases} \quad \text{for } a, b \geq 0, \tag{17}$$

and (10), as well as the abbreviations

$$d_{n,k} = D_p^{x_0}(x^\dagger, z_{n,k})^{1/2} \quad (18)$$

$$t_{n,k} = \|A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta\| \quad (19)$$

$$\tilde{t}_{n,k} = \|A_n^* j_r^Y(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta)\| \quad (20)$$

$$\leq \|A_n\| t_{n,k}^{-1}. \quad (21)$$

$$r_n = \|F(x_n^\delta) - y^\delta\| \quad (22)$$

Here

$$\varphi(\lambda) = 2^{s^*-1} C_{p^*, s^*} (p\rho^2)^{1-\frac{s^*}{p^*}} \lambda^{s^*} + 2^{p^*-1} C_{p^*, s^*} \lambda^{p^*}, \quad (23)$$

which by $p^* \geq s^* > 1$ defines a strictly monotonically increasing and convex function on \mathbb{R}^+ . Note that estimate (16) is just the same as (15) in [9].

For the term (II) in (14) we get, using (9), (2),

$$\begin{aligned} & \omega_{n,k} \langle j_r^Y(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta), A_n(z_{n,k} - x^\dagger) \rangle_{Y^* \times Y} \\ &= \omega_{n,k} t_{n,k}^r \\ & \quad + \omega_{n,k} \langle j_r^Y(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta), \\ & \quad \quad A_n(x_n^\delta - x^\dagger) - F(x_n^\delta) + y^\delta \rangle_{Y^* \times Y} \\ & \geq \omega_{n,k} t_{n,k}^r - \omega_{n,k} t_{n,k}^{r-1} (\eta \|F(x_n^\delta) - y^\delta\| + (1+\eta)\delta) \end{aligned} \quad (24)$$

This is the same as (19) in [9].

To make use of the variational inequality (11) for estimating (III), we first of all use (9) to conclude

$$\begin{aligned} & \|F(z_{n,k}) - F(x^\dagger)\| \\ &= \|(A_n(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta) \\ & \quad + (F(z_{n,k}) - F(x_n^\delta) - A_n((z_{n,k} - x_n^\delta) + (y^\delta - y))\| \\ &\leq t_{n,k} + \eta \|F(z_{n,k}) - F(x_n^\delta)\| + \delta \\ &\leq t_{n,k} + \eta (\|F(z_{n,k}) - F(x^\dagger)\| + \|F(x_n^\delta) - y^\delta\|) + (1+\eta)\delta, \end{aligned}$$

hence

$$\|F(z_{n,k}) - F(x^\dagger)\| \leq \frac{1}{1-\eta} (t_{n,k} + \eta r_n + (1+\eta)\delta). \quad (25)$$

This together with (11) implies

$$\begin{aligned} & |\alpha_{n,k} \langle J_p^X(x^\dagger - x_0), z_{n,k} - x^\dagger \rangle_{X^* \times X}| \\ &\leq \frac{\beta}{(1-\eta)^{2\nu}} \alpha_{n,k} d_{n,k}^{1-2\nu} (t_{n,k} + \eta r_n + (1+\eta)\delta)^{2\nu} \\ &\leq C(\lambda(s, \nu)) \frac{\beta}{(1-\eta)^{2\nu}} \alpha_{n,k} \left\{ d_{n,k}^2 + (t_{n,k} + \eta r_n + (1+\eta)\delta)^{\frac{4\nu}{1+2\nu}} \right\} \end{aligned} \quad (26)$$

where we have used the elementary estimate

$$a^{1-\lambda} b^\lambda \leq C(\lambda)(a+b) \quad \text{for } a, b \geq 0, \lambda \in (0, 1) \quad (27)$$

with $C(\lambda) = \lambda^\lambda (1-\lambda)^{1-\lambda}$ and $\lambda(s, \nu)$ as in (12). Note that (26) differs from the corresponding estimate (24) in [9].

Finally, for the term (IV) we have that

$$\begin{aligned} & \alpha_{n,k} \langle J_p^X(x^\dagger - x_0) - J_p^X(z_{n,k} - x_0), x^\dagger - z_{n,k} \rangle_{X^* \times X} \\ &= \alpha_{n,k} (D_p^{x_0}(x^\dagger, z_{n,k}) + D_p^{x_0}(z_{n,k}, x^\dagger)) \geq \alpha_{n,k} d_{n,k}^2, \end{aligned} \quad (28)$$

which is just (26) in [9].

Altogether we arrive at the estimate

$$\begin{aligned} d_{n,k+1}^2 &\leq \left(1 - (1 - c_0)\alpha_{n,k}\right) d_{n,k}^2 + c_1 \alpha_{n,k}^{s^*} + c_2 \alpha_{n,k}^{p^*} \\ &\quad + c_3 \alpha_{n,k} (t_{n,k} + \eta r_n + (1 + \eta)\delta)^{\frac{4\nu}{1+2\nu}} \\ &\quad - \omega_{n,k} t_{n,k}^r + \omega_{n,k} t_{n,k}^{r-1} (\eta r_n + (1 + \eta)\delta) + \varphi(\omega_{n,k} \tilde{t}_{n,k}), \end{aligned} \quad (29)$$

where

$$c_0 = \frac{\beta}{(1 - \eta)^{2\nu}} C(\lambda(s, \nu)) \quad (30)$$

$$c_1 = C_{p^*, s^*} (p\rho^2)^{1 - \frac{s^*}{p^*}} \bar{\rho}^{(p-1)s^*} 2^{s^*-1} \quad (31)$$

$$c_2 = C_{p^*, s^*} \bar{\rho}^p 2^{p^*-1} \quad (32)$$

$$c_3 = \frac{\beta}{(1 - \eta)^{2\nu}} C(\lambda(s, \nu)) \quad (33)$$

$$\theta = \frac{4\nu}{r(1 + 2\nu) - 4\nu}, \quad (34)$$

(small c denoting constants that can be made small by assuming x_0 to be sufficiently close to x^\dagger and therewith $\beta, \eta, \|x_0 - x^\dagger\|$ small), and $\lambda(s, \nu)$ as in (12).

Multiplying (29) with $\alpha_{n,k+1}^{-\theta}$ and abbreviating

$$\gamma_{n,k} := d_{n,k}^2 \alpha_{n,k}^{-\theta},$$

we get

$$\begin{aligned} &\gamma_{n,k+1} - \gamma_{n,k} \\ &\leq \left(\frac{\alpha_{n,k}}{\alpha_{n,k+1}}\right)^\theta \left\{ \left(1 - (1 - c_0)\alpha_{n,k} - \left(\frac{\alpha_{n,k+1}}{\alpha_{n,k}}\right)^\theta\right) \gamma_{n,k} \right. \\ &\quad + (c_1 \alpha_{n,k}^{s^*-\theta} + c_2 \alpha_{n,k}^{p^*-\theta} + c_3 \alpha_{n,k}^{1-\theta} (t_{n,k} + \eta r_n + (1 + \eta)\delta)^{\frac{4\nu}{1+2\nu}} \\ &\quad \left. - \alpha_{n,k}^{-\theta} (\omega_{n,k} t_{n,k}^r - \omega_{n,k} t_{n,k}^{r-1} (\eta r_n + (1 + \eta)\delta) - \varphi(\omega_{n,k} \tilde{t}_{n,k})) \right\}. \end{aligned} \quad (35)$$

To obtain monotone decay of the sequence $\gamma_{n,k}$ with increasing k we choose

- $\omega_{n,k} \geq 0$ such that

$$\underline{\omega} \leq \omega_{n,k} \leq \bar{\omega} \text{ and } \frac{\varphi(\omega_{n,k} \tilde{t}_{n,k})}{\omega_{n,k} t_{n,k}^r} \leq \bar{c}_\omega \quad (36)$$

for some $0 < \underline{\omega} < \bar{\omega}$, $\bar{c}_\omega > 0$. We will do so by setting

$$\omega_{n,k} = \vartheta \min\{t_{n,k}^{\frac{r}{s^*-1}} \tilde{t}_{n,k}^{-s}, t_{n,k}^{\frac{r}{p^*-1}} \tilde{t}_{n,k}^{-p}, \bar{\omega}\} \quad (37)$$

with ϑ sufficiently small, cf. [9], and assuming that

$$r \geq s \geq p, \quad (38)$$

- $\alpha_{n,k} \geq 0$ such that

$$\alpha_{n,k} \geq \check{\alpha}_{n,k} := \tilde{\tau} (t_{n,k} + \eta r_n + (1 + \eta)\delta)^{\frac{r}{1+\theta}} \quad (39)$$

and

$$c_0 + \frac{c_1}{\bar{\gamma}_{0,0}} \alpha_{n,k}^{s^*-\theta-1} + \frac{c_2}{\bar{\gamma}_{0,0}} \alpha_{n,k}^{p^*-\theta-1} + \frac{c_3}{\tilde{\tau}^\theta \bar{\gamma}_{0,0}} + \frac{1}{\tilde{\tau}^{1+\theta} \bar{\gamma}_{0,0}} \leq q < 1. \quad (40)$$

The latter can be achieved by

$$\alpha_{n,k} \leq 1 \text{ and} \quad (41)$$

$$s^* \geq \theta + 1, \quad p^* \geq \theta + 1, \quad (42)$$

$c_0, c_1, c_2, c_3, \tilde{\tau}$ sufficiently small.

In case $\nu > 0$ we additionally require

$$\alpha_{n,k+1} \geq \hat{\alpha}_{n,k+1} := \alpha_{n,k} \left(1 - (1-q)\alpha_{n,k}\right)^{1/\theta} \quad (43)$$

with an upper bound $\bar{\gamma}_{0,0}$ for $\gamma_{0,0}$. Note that this just means $\alpha_{n,k+1} \geq 0$ in case $\nu = 0$, i.e., $\theta = 0$, thus an empty condition in this case.

To meet conditions (39), (43) with a minimal $\alpha_{n,k+1}$ we set

$$\begin{aligned} \alpha_{n,k+1} &= \max\{\tilde{\alpha}_{n,k+1}, \hat{\alpha}_{n,k+1}\} \text{ for } k \geq 0 \\ \alpha_{n,0} &= \begin{cases} \alpha_{n-1,k_{n-1}} & \text{if } n \geq 1 \\ \alpha_{0,0} & \text{if } n = 0 \end{cases} \end{aligned} \quad (44)$$

It remains to choose

- the inner stopping index k_n and
- the outer stopping index n_* ,

see below.

Indeed with these choices of $\omega_{n,k}$ and $\alpha_{n,k+1}$ we can inductively conclude from (35) that

$$\begin{aligned} \gamma_{n,k+1} - \gamma_{n,k} &\leq \left(\frac{\alpha_{n,k}}{\alpha_{n,k+1}}\right)^\theta \left\{ \left(1 - (1-q)\alpha_{n,k} - \left(\frac{\alpha_{n,k+1}}{\alpha_{n,k}}\right)^\theta\right) \bar{\gamma}_{0,0} \right\} \\ &\quad - \alpha_{n,k+1}^{-\theta} (1 - \bar{c}_\omega) \omega_{n,k} t_{n,k}^r, \\ &\leq -\alpha_{n,k+1}^{-\theta} (1 - \bar{c}_\omega) \omega_{n,k} t_{n,k}^r \leq 0. \end{aligned} \quad (45)$$

This monotonicity result holds for all $n \in \mathbb{N}$ and for all $k \in \mathbb{N}$.

By (45) and $\alpha_{n,k} \leq 1$ (cf. (41)) it can be shown inductively that all iterates remain in $\mathcal{B}_\rho^D(x^\dagger)$ provided

$$\bar{\gamma}_{0,0} \leq \rho^2. \quad (46)$$

Moreover, (45) implies that

$$\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \alpha_{n,k+1}^{-\theta} \omega_{n,k} t_{n,k}^r \leq \frac{\bar{\gamma}_{0,0}}{1 - \bar{c}_\omega} < \infty, \quad (47)$$

hence by $\alpha_{n,k+1} \leq 1$, $\omega_{n,k} \geq \omega$

$$t_{n,k} \rightarrow 0 \text{ as } k \rightarrow \infty \text{ for all } n \in \mathbb{N} \quad (48)$$

and

$$\sup_{k \in \mathbb{N}_0} t_{n,k} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (49)$$

Especially, since $t_{n,0} = r_n$, (47) implies

$$\sum_{n=0}^{\infty} r_n^r < \infty, \quad (50)$$

hence $r_n \rightarrow 0$ as $n \rightarrow \infty$.

To quantify the behavior as $k \rightarrow \infty$ of the sequence $\alpha_{n,k}$ according to (39), (43), (44) for fixed n we distinguish between two cases.

Case (i): there exists a \underline{k} such that for all $k \geq \underline{k}$ we have $\alpha_{n,k} = \hat{\alpha}_{n,k}$. Considering

an arbitrary accumulation point $\bar{\alpha}_n$ of $\alpha_{n,k}$ (which exists since $0 \leq \alpha_{n,k} \leq 1$) we therefore have $\bar{\alpha}_n = \bar{\alpha}_n \left(1 - (1 - q)\bar{\alpha}_n\right)^{\frac{1}{\theta}}$, hence $\bar{\alpha}_n = 0$.

Case (ii): consider the situation that (i) does not hold, i.e., there exists a subsequence k_j such that for all $j \in \mathbb{N}$ we have $\alpha_{n,k_j} = \check{\alpha}_{n,k_j}$. Then by (39), (43), and (48) we have $\alpha_{n,k_j} \rightarrow \tilde{\tau}(\eta r_n + (1 + \eta)\delta)^{\frac{r}{1+\theta}}$. Altogether we have shown that

$$\limsup_{k \rightarrow \infty} \alpha_{n,k} \leq \tilde{\tau}(\eta r_n + (1 + \eta)\delta)^{\frac{r}{1+\theta}} \text{ for all } n \in \mathbb{N}. \quad (51)$$

Since η and δ can be assumed to be sufficiently small, this especially implies the bound $\alpha_{n,k} \leq 1$ in (41).

We consider $z_{n_*, k_{n_*}^*}$ as our regularized solution, where n_* , $k_{n_*}^*$ (and also k_n for all $n \leq n_* - 1$; note that $k_{n_*}^*$ is to be distinguished from k_{n_*} - actually the latter is not defined, since we only define k_n for $n \leq n_* - 1$!) are still to be chosen appropriately, according to the requirements from the proofs of

- convergence rates in case $\nu, \theta > 0$,
- convergence for exact data $\delta = 0$,
- convergence for noisy data as $\delta \rightarrow 0$.

3.1. Convergence rates in case $\nu, \theta > 0$. From (45) we get

$$d_{n,k}^2 \leq \bar{\gamma}_{0,0} \alpha_{n,k}^\theta \text{ for all } n, k \in \mathbb{N}, \quad (52)$$

hence in order to get the desired rate

$$d_{n_*, k_{n_*}^*}^2 = O(\delta^{\frac{r\theta}{1+\theta}})$$

in view of (51) (which is a sharp bound in case (ii) above) we need to have a bound

$$r_{n_*} \leq \tau \delta \quad (53)$$

for some constant $\tau > 0$, and we should choose $k_{n_*}^*$ large enough so that

$$\alpha_{n_*, k_{n_*}^*} \leq C_\alpha (r_{n_*} + \delta)^{\frac{r}{1+\theta}} \quad (54)$$

which is possible with a finite $k_{n_*}^*$ by (51) for $C_\alpha > (\tilde{\tau}(1 + \eta))^{\frac{r}{1+\theta}}$. Note that this holds without any requirements on k_n for $n < n_*$.

3.2. Convergence as $n \rightarrow \infty$ for exact data $\delta = 0$. To show that $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence (following the seminal paper [4]), for arbitrary $m < j$, we choose the index $l \in \{m, \dots, j\}$ such that r_l is minimal and use the identity

$$\begin{aligned} D_p^{x_0}(x_l, x_m) &= D_p^{x_0}(x^\dagger, x_m) - D_p^{x_0}(x^\dagger, x_l) \\ &\quad + \langle J_p^X(x_l - x_0) - J_p^X(x_m - x_0), x_l - x^\dagger \rangle_{X^* \times X} \end{aligned} \quad (55)$$

and the fact that the monotone decrease and boundedness from below of the sequence $D_p^{x_0}(x^\dagger, x_m)$ implies its convergence, hence it suffices to prove that the last term in (55) tends to zero as $m < l \rightarrow \infty$. (Analogously it can be shown that $D_p^{x_0}(x_l, x_j)$ tends to zero as $l < j \rightarrow \infty$). This term can be rewritten as

$$\begin{aligned} &\langle J_p^X(x_l - x_0) - J_p^X(x_m - x_0), x_l - x^\dagger \rangle_{X^* \times X} \\ &= \sum_{n=m}^{l-1} \sum_{k=0}^{k_n-1} \langle u_{n,k+1} - u_{n,k}, x_l - x^\dagger \rangle_{X^* \times X}, \end{aligned}$$

where

$$\begin{aligned}
& |\langle u_{n,k+1} - u_{n,k}, x_l - x^\dagger \rangle_{X^* \times X}| \\
&= |\alpha_{n,k} \langle J_p^X(z_{n,k} - x_0), x_l - x^\dagger \rangle_{X^* \times X} \\
&\quad + \omega_{n,k} \langle j_r^Y(A_n(z_{n,k} - x_n^\delta) - b_n), A_n(x_l - x^\dagger) \rangle_{X^* \times X}| \\
&\leq 2\bar{\rho}^p \alpha_{n,k} + \omega_{n,k} t_{n,k}^{r-1} \|A_n(x_l - x^\dagger)\| \\
&\leq 2\bar{\rho}^p \tilde{\tau} (t_{n,k} + \eta r_n)^r + \omega_{n,k} t_{n,k}^{r-1} (1 + \eta) (2r_n + r_l) \\
&\leq 2\bar{\rho}^p \tilde{\tau} (t_{n,k} + \eta r_n)^r + 3(1 + \eta) \omega_{n,k} t_{n,k}^{r-1} r_n
\end{aligned}$$

by our choice of $\alpha_{n,k} = \check{\alpha}_{n,k}$ (note that $\hat{\alpha}_{n,k} = 0$ in case $\theta = 0$), condition (9) and minimality of r_l . Thus we have by $\omega_{n,k} \leq \bar{\omega}$ and Young's inequality that there exists $C > 0$ such that

$$\langle J_p^X(x_l - x_0) - J_p^X(x_m - x_0), x_l - x^\dagger \rangle_{X^* \times X} \leq C \sum_{n=m}^{l-1} \left\{ \left(\sum_{k=0}^{k_n-1} t_{n,k}^r \right) + k_n r_n^r \right\}$$

for which we can conclude convergence as $m, l \rightarrow \infty$ from (47) provided that

$$\sum_{n=m}^{\infty} k_n r_n^r \rightarrow 0 \text{ as } m \rightarrow \infty,$$

which we guarantee by choosing, for an a priori fixed summable sequence $(a_n)_{n \in \mathbb{N}}$, e.g. $a_n = 2^{-n}$

$$k_n := [a_n r_n^{-r}], \quad (56)$$

or, using (50) just $k_n \equiv \bar{k}$ for some fixed integer \bar{k} , e.g., $\bar{k} = 3$. This is consistent with (54), since in case $\delta = 0$ we have $n_* = \infty$, so condition (54) never gets active in the noiseless case.

3.3. Convergence with noisy data as $\delta \rightarrow 0$. In case $\nu, \theta > 0$, convergence follows from the convergence rates results in Subsection 3.1. Therefore it only remains to show convergence as $\delta \rightarrow 0$ in case $\nu, \theta = 0$.

In this section we explicitly emphasize dependence of the computed quantities on the noisy data and on the noise level by a superscript δ .

Let $\|y^{\delta_j} - y\| \leq \delta_j$ with δ_j a zero sequence and n_{*j} the corresponding stopping index. As usual [4] we distinguish between the two cases that (i) n_{*j} has a finite accumulation point and (ii) n_{*j} tends to infinity.

Case (i): there exists an $N \in \mathbb{N}$ and a subsequence n_{j_i} such that for all $i \in \mathbb{N}$ we have $n_{j_i} = N$. Provided

$$n_*(\delta) = N \text{ for all } \delta \Rightarrow \text{The mapping } \delta \mapsto x_N^\delta \text{ is continuous at } \delta = 0, \quad (57)$$

we can conclude that $x_N^{\delta_{j_i}} \rightarrow x_N^0$ as $i \rightarrow \infty$, and by taking the limit as $i \rightarrow \infty$ also in (53), x_N^0 is a solution to (1). Thus we may set $x^\dagger = x_N^0$ in (45) (with $\theta = 0$) to obtain

$$D_p^{x_0}(x_N^0, z_{n_{*j_i}, k_{n_{*j_i}}}^{\delta_{j_i}}) = D_p^{x_0}(x_N^0, z_{N, k_{n_{*j_i}}}^{\delta_{j_i}}) \leq D_p^{x_0}(x_N^0, x_N^{\delta_{j_i}}) \rightarrow 0 \text{ as } i \rightarrow \infty,$$

where we have again used the continuous dependence (57) in the last step.

Case (ii): let $n_{*j} \rightarrow \infty$ as $j \rightarrow \infty$, and let x^\dagger be a solution to (1). For arbitrary $\epsilon > 0$, by convergence for $\delta = 0$ (see the previous subsection) we can find n such

that $D_p^{x_0}(x^\dagger, x_n^0) < \frac{\epsilon}{2}$ and, by Theorem 2.60 (d) in [21] there exists j_0 such that for all $j \geq j_0$ we have $n_{*,j} \geq n+1$ and $|D_p^{x_0}(x^\dagger, x_n^{\delta_j}) - D_p^{x_0}(x^\dagger, x_n^0)| < \frac{\epsilon}{2}$, provided

$$n \leq n_*(\delta) - 1 \text{ for all } \delta \Rightarrow \text{The mapping } \delta \mapsto x_n^\delta \text{ is continuous at } \delta = 0. \quad (58)$$

Hence, by monotonicity of the errors we have

$$D_p^{x_0}(x^\dagger, z_{n_{*,j}, k_{n_{*,j}}}^{\delta_j}) \leq D_p^{x_0}(x^\dagger, x_n^{\delta_j}) \leq D_p^{x_0}(x^\dagger, x_n^0) + |D_p^{x_0}(x^\dagger, x_n^{\delta_j}) - D_p^{x_0}(x^\dagger, x_n^0)| < \epsilon.$$

Indeed, (57), (58) can be concluded from continuity of F, F' , the definition of the method (3), as well as stable dependence of all parameters $\omega_{n,k}, \alpha_{n,k}, k_n$ according to (37), (39), (43), (44), (56) on the data y^δ .

Altogether we have derived the following algorithm.

Algorithm 3.1. (Newton – iteratively regularized Landweber method)

Choose $\tau, \tilde{\tau}, C_\alpha$ sufficiently large, x_0 sufficiently close to x^\dagger ,
 $\alpha_{00} \leq 1, \vartheta > 0$ sufficiently small, $\bar{\omega} > 0, (a_n)_{n \in \mathbb{N}_0}$ such that $\sum_{n=0}^\infty a_n < \infty$
 For $n = 0, 1, 2, \dots$ until $r_n \leq \tau\delta$ do
 $u_{n,0} = 0$
 $z_{n,0} = x_n^\delta$
 $\alpha_{n,0} = \alpha_{n-1, k_{n-1}}$ if $n > 0$
 For $k = 0, 1, 2, \dots$ until $\left\{ \begin{array}{ll} k = k_n - 1 = a_n r_n^{-r} & \text{if } r_n > \tau\delta \\ \alpha_{n,*, k_{n,*,*}} \leq C_\alpha (r_{n,*} + \delta)^{\frac{r}{1+\vartheta}} & \text{if } r_n \leq \tau\delta \end{array} \right\}$ do
 $\omega_{n,k} = \vartheta \min\{\tilde{t}_{n,k}^{\frac{r}{s^*-1}} \tilde{t}_{n,k}^{-s}, \tilde{t}_{n,k}^{\frac{r}{p^*-1}} \tilde{t}_{n,k}^{-p}, \bar{\omega}\}$
 $u_{n,k+1} = u_{n,k} - \alpha_{n,k} J_p^X(z_{n,k} - x_0) - \omega_{n,k} F'(x_n^\delta)^* J_r^Y(F'(x_n^\delta)(z_{n,k} - x_n^\delta) + F(x_n^\delta) - y^\delta)$
 $z_{n,k+1} = x_0 + J_p^{X^{-1}}(J_p^X(x_n^\delta - x_0) + u_{n,k+1})$
 $\alpha_{n,k+1} = \max\{\check{\alpha}_{n,k+1}, \hat{\alpha}_{n,k+1}\}$ with $\check{\alpha}_{n,k+1}, \hat{\alpha}_{n,k+1}$ as in (39), (43)
 $x_{n+1}^\delta = z_{n,k_n}.$

Here we use the abbreviations according to (15), (19), (20), (22), (34).

The analysis above yields the following convergence result.

Theorem 3.2. Assume that X is smooth and s -convex with $s \geq p$, that x_0 is sufficiently close to x^\dagger , i.e., $x_0 \in \mathcal{B}_\rho^D(x^\dagger)$, that F satisfies (9) with (8), that F and F' are continuous and uniformly bounded in $\mathcal{B}_\rho^D(x^\dagger)$, and that (38), (42) hold. Then, the iterates $z_{n,k}$ defined by Algorithm 3.1 remain in $\mathcal{B}_\rho^D(x^\dagger)$ and converge to a solution x^\dagger of (1) subsequentially as $\delta \rightarrow 0$ (i.e., there exists a convergent subsequence and the limit of every convergent subsequence is a solution). In case of exact data $\delta = 0$, we have subsequential convergence of x_n to a solution of (1) as $n \rightarrow \infty$.

If additionally a variational inequality (11) with $\nu \in (0, 1]$ and β sufficiently small is satisfied, we obtain optimal convergence rates

$$D_p^{x_0}(x^\dagger, z_{n_{*,k_{n_{*,*}}}}) = O(\delta^{\frac{4\nu}{2\nu+1}}), \quad \text{as } \delta \rightarrow 0. \quad (59)$$

Note that we here deal with an a priori parameter choice: θ and therefore ν has to be known, otherwise ν must be set to a lower bound $\underline{\nu}$ for the true ν and since $\underline{\nu} \leq \nu$ implies validity of (11) with ν replaced by $\underline{\nu}$, Theorem 3.2 still implies the (possibly suboptimal) rates $O(\delta^{\frac{4\underline{\nu}}{2\underline{\nu}+1}})$, or just convergence if we have set $\underline{\nu} = 0$.

4. Numerical Experiments. In this section we present some numerical experiments to test the method defined in section 3. We consider the identification of the space-dependent coefficient c in the elliptic boundary value problem

$$\begin{cases} -\Delta u + cu = f, & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \quad (60)$$

from the measurement of u in Ω , where f is a fixed function and where Ω is assumed to be a smooth, bounded domain in \mathbb{R}^d , $d \in \mathbb{N}$. Note that inhomogeneous Dirichlet boundary conditions can be easily incorporated into the right-hand side f if necessary.

We consider three examples with $d = 1$ and an example with $d = 2$. In all cases, we take $X := L^p(\Omega)$, $1 < p < \infty$, $Y := L^r(\Omega)$, $1 < r < \infty$ and recall that the following facts hold true.

- (i) For $1 < p < \infty$, the duality mapping in X is given by $J_p^X(c) = |c|^{p-1} \text{sgn}(c)$.
- (ii) If the domain of the forward operator is defined by

$$\mathcal{D}(F) := \{c \in X \mid \|c - \hat{c}\|_X \leq \gamma_p, \text{ for some } \hat{c} \in L^\infty(\Omega), \hat{c} \geq 0 \text{ a.e.}\}, \quad (61)$$

the condition (8) is satisfied with $c_0 = 0$.

- (iii) There follows from Lemma 2 in [15] that the operator $F : \mathcal{D}(F) \subseteq X \rightarrow Y$, $F(c) = A(c)^{-1}f$ is well defined. Here $A(c) : W^{2,p}(\Omega) \cap H_0^1(\Omega) \rightarrow L^p(\Omega)$ is given by $A(c)u := -\Delta u + cu$. Moreover, $F'(c)$ and the adjoint of $F'(c)$

$$F'(c)h = -A(c)^{-1}(hF(c)), \quad F'(c)^*w = -u(c)A(c)^{-1}w,$$

are well defined and bounded.

In all the numerical simulations, we take $\theta = 0$ and stop the outer iteration by means of the discrepancy principle (53). Concerning the stopping index of the inner iteration, we slightly modify Algorithm 1, requiring also that if $\|F(z_{n,k}) - y^\delta\| \leq \tau\delta$ then the iteration has to be stopped. More precisely,

$$k_n = \min\{k \in \mathbb{Z}, k \geq 0, \mid \|F(z_{n,k}) - y^\delta\| \leq \tau\delta \vee k \geq a_n r_n^{-r}\} \quad (62)$$

and the regularized solution is $c_{n_*}^\delta = z_{n_{n_*-1}, k_{n_*-1}}$.

4.1. 1-dimensional examples. We consider the same numerical simulations as in [9], taking $\Omega = (0, 1)$ and inhomogeneous boundary conditions $u(0) = g_0$, $u(1) = g_1$. We solve all differential equations approximately by a finite difference method by dividing the interval $[0, 1]$ into $N + 1$ subintervals with equal length $1/(N + 1)$, in all examples below $N = 400$. The L^p and L^r norms are calculated approximately by means of a quadrature method.

Example 4.1. In the first simulation we assume that the solution is sparse:

$$c^\dagger(t) = \begin{cases} 0.5, & 0.3 \leq t \leq 0.4, \\ 1.0, & 0.6 \leq t \leq 0.7, \\ 0.0, & \text{elsewhere.} \end{cases} \quad (63)$$

The test problem is constructed by taking $u(t) = u(c^\dagger)(t) = 1 + 5t$, $f(t) = u(t)c^\dagger(t)$, $g_0 = 1$ and $g_1 = 6$. We perturb the exact data u with gaussian white noise: the corresponding perturbed data u^δ satisfies $\|u^\delta - u\|_{L^r} = \delta$, with $\delta = 0.1 \times 10^{-3}$.

We apply Algorithm 1, with the inner stopping index satisfying (62), with $\tau = 1.02$, $\tilde{\tau} = 0.1$, $a_n = (50 + n)^{-2}$. The upper bound \bar{c}_ω is fixed equal to 0.1 and ϑ is chosen as 2^{-j^\sharp} , where j^\sharp is the first index that satisfies

$$2^{s^*-1}C_{p^*,s^*}(p\rho^2)^{1-s^*/p^*}\vartheta^{s^*-1} + 2^{p^*-1}C_{p^*,s^*}\vartheta^{p^*-1} \leq \bar{c}_\omega. \quad (64)$$

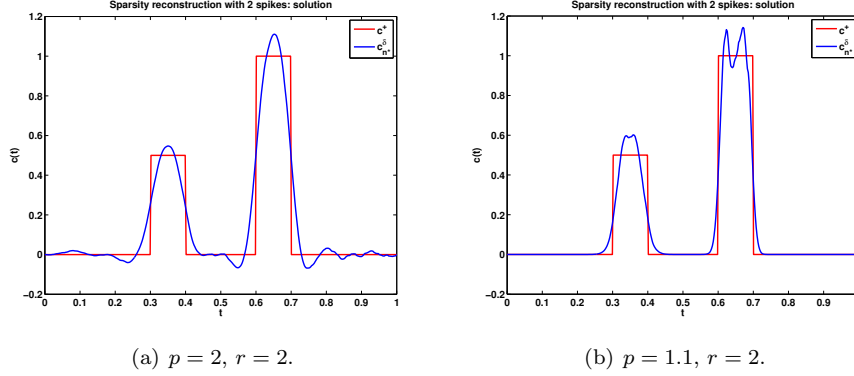


FIGURE 1. Reconstructed Solutions for example 4.1

In figure 1 we show the results obtained by our method with $p = 2$ and $p = 1.1$ respectively. The reconstructed solutions are very similar to those obtained in [9] and [10]. Concerning the total number of inner iterations

$$N_p = \sum_{n=0}^{n_*-1} k_n,$$

similarly to [9], it is larger in the case $p = 2$ ($N_2 = 3992$) than in the case $p = 1.1$ ($N_{1.1} = 3063$). In both cases, the value of N_p is lower than the corresponding value found in [9]. We underline that it is possible to make different choices for $\tilde{\tau}$, \bar{c}_ω , a_n to look for further improvements of the speed of the method. The partial freedom in the choices of these parameters makes Algorithm 1 more flexible than the method described in [9]. In our numerical simulations, we tested different choices which gave similar but slightly worse results than those stated here.

Example 4.2. We modify the exact solution of the previous example into:

$$c^\dagger(t) = \begin{cases} 0.25, & 0.1 \leq t \leq 0.15, \\ 0.5, & 0.3 \leq t \leq 0.4, \\ 1.0, & 0.6 \leq t \leq 0.7, \\ 0.0 & \text{elsewhere.} \end{cases} \quad (65)$$

and choose again $\delta = 0.1 \times 10^{-3}$. In this case, we take $\tau = 1.02$, $\tilde{\tau} = 0.01$, $a_n = (100 + n)^{-2}$, $\bar{c}_\omega = 0.1$ and ϑ as in the previous example. In figure 2 we show the results obtained by our method with $p = 2$ and $p = 1.1$ respectively. As usual, the reconstruction of the sparsity is much better for $p = 1.1$. Concerning the total number of inner iteration N_p , in this case we obtain $N_2 = 4141$ (with a corresponding error of 0.1110) and $N_{1.1} = 3110$ (with a corresponding error of 0.0482). We observe that although the corresponding error is slightly larger than that obtained in [9], the value of N_2 is slightly more than a fifth than the value obtained in [9], with a gain in the speed of the 421.8%. Moreover, in the case $p = 1.1$ Algorithm 1 performs 1415 iterations less than in [9], with a gain in the speed of the 45.5%, and obtains even a lower error.

Example 4.3. We consider an example with noisy data where a few data points called outliers are remarkably different from other data points. This situation may

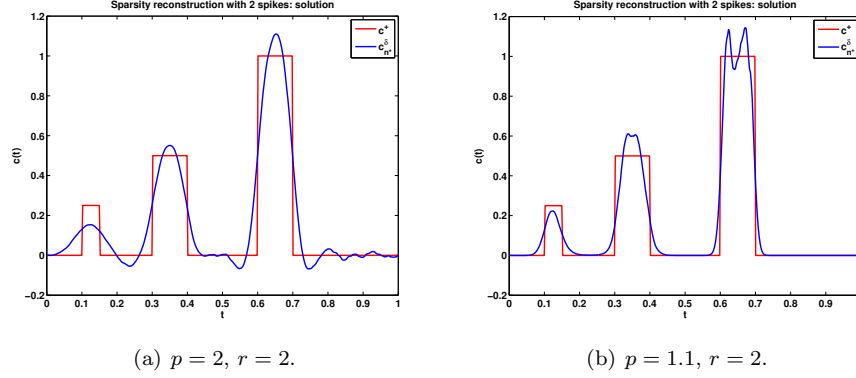


FIGURE 2. Reconstructed Solutions for example 4.2

arise from procedural measurement errors.

We suppose c^\dagger to be a smooth solution

$$c^\dagger(t) = 2 - t + 4 \sin(2\pi t) \quad (66)$$

and take $u(c^\dagger)(t) = 1 - 2t$, $f(t) = (1 - 2t)(2 - t + 4 \sin(2\pi t))$, $u(0) = g_0 = 1$ and $u(1) = g_1 = -1$ as exact data of the problem. We start the iteration from the initial guess $c_0(t) = 2 - t$, and take $\tilde{\tau} = 5 \times 10^{-3}$, $a_n = (1 + n)^{-1.1}$, $\bar{c}_\omega = 5 \times 10^{-3}$ and ϑ as in the previous example. Using the same Matlab seed for generating random data, we consider the same perturbed data as in [9], Example 3, case *B* (cf. Figure 3 here and Figure 3, picture (d) in that paper).

We run both Algorithm 1 with $\tilde{\tau} = 1.0015$ and the algorithm that generated the results presented in [9] for this example with $p = 2$ and $r = 1.1$ ¹. Pictures (b) and (d) from figure 3 show the corresponding results. The solution obtained by Algorithm 1 is slightly more precise, with an error equal to 1.8852×10^{-1} for Algorithm 1 and equal to 2.9388×10^{-1} for the method in [9]. The most interesting fact is that Algorithm 1 computes only $N_2 = 249$ total inner iterations to obtain this solution, whereas for the method in [9] $N_2 = 3285$ and the reconstruction is poorer. Moreover, due to the flexibility of our method, we can simply change the value of τ into $1 + 10^{-5}$ to get a more precise solution (see picture (c) in figure 3). A visual inspection gives an idea of the improvement: the error of this solution, obtained with 278 total inner iterations, is equal to 1.1607×10^{-1} .

We summarize the numerical results of the 1-dimensional examples in Table 1.

4.2. A 2-dimensional example. We show the performance of Algorithm 1 in the following 2-dimensional example.

Let $\Omega = (0, 1)^2$ and assume the exact solution to be

$$c^\dagger(x, y) = 40\chi_{[0.19, 0.24]^2}(x, y), \quad x, y \in \Omega, \quad (67)$$

where the function χ is the characteristic function on a subset of \mathbb{R}^2 . We take as exact data $u(x, y) = 1 + x + y$: as a consequence, the fixed right hand side of the problem is $f(x, y) = c^\dagger(x, y)u(x, y)$ and the data at the boundary are given

¹These values of r and p do not satisfy the condition (38) of Theorem 3.2. However, this condition is needed only to have a lower bound for $\omega_{n,k}$ and such a bound is verified experimentally in this example for these values of r and p .

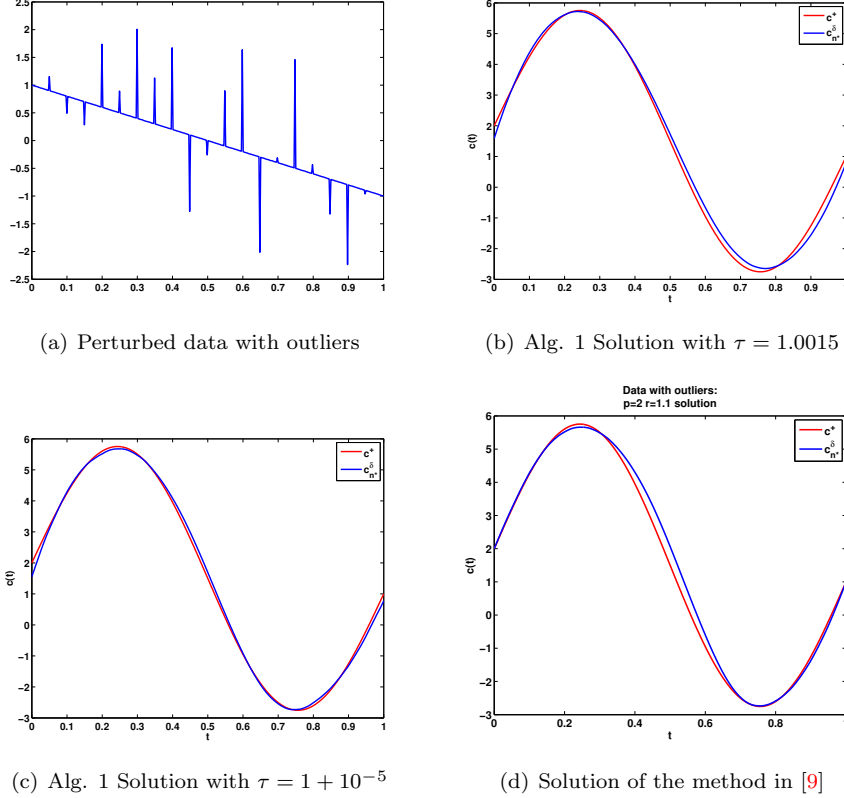


FIGURE 3. Numerical results for example 4.3: (a) are the perturbed data; (b) is the solution obtained by the method in [9]; (c) and (d) are the solutions of Algorithm 1 with different values of τ .

by $g(x, y) = u(x, y)$ for every $(x, y) \in \partial \Omega$. We discretize the interval $[0, 1]_x$ into $N + 1$ subintervals and the interval $[0, 1]_y$ into $M + 1$ subintervals and compute the solutions of the forward operator $F(c) = u(c)$ by a finite difference method. The L^p and L^r norms in Ω are calculated by a simple quadrature method.

We fix $p = 1.1$, $N = M = 30$, $\tau = 1 + 10^{-5}$, $\tilde{\tau} = 10^{-4}$, $\bar{c}_\omega = 0.1$ and ϑ as in the 1-dimensional examples. We run Algorithm 1 starting from $c_0 = 0$ with the data u^δ perturbed by white gaussian noise with two different noise levels. In the first case, we choose a small value for $\delta = \|u - u^\delta\| = 10^{-3}$ with $r = 2$. In the second case we choose $\delta = 10^{-2}$ with $r = 2$ and with $r = 10$.

In figure 4 we plot the exact solution and the reconstructions obtained using Algorithm 1 in all these cases. The pictures show that the method provides a good reconstruction of the sparsity in this example. In particular, we underline that in the case $\delta = 10^{-2}$ the choice of a large r improves the result, obtaining a better reconstruction with very few iterations (the total number of iterations is equal to 9 in this case).

5. Conclusions. In this paper we have devised an alternative parameter choice strategy for the iteratively regularized Newton- Landweber iteration proposed in

1-dimensional numerical simulations		
Example 1		
Method	Total n. of inner iterations	Error: $\ \cdot - c^\dagger\ $
Alg. 1 $p = 1.1$	3063	0.0413
Alg. 1 $p = 2.0$	3992	0.1059
Example 2		
Method	Total n. of inner iterations	Error: $\ \cdot - c^\dagger\ $
Alg. 1 $p = 1.1$	3110	0.0482
Alg. 1 $p = 2.0$	4141	0.1110
Example 3		
Method	Total n. of inner iterations	Error: $\ \cdot - c^\dagger\ $
Alg. 1 $\tau = 1.0015$	249	0.1885
Alg. 1 $\tau = 1 + 10^{-5}$	278	0.1161
Method from [9]	3285	0.2939

TABLE 1. Numerical results for the 1-dimensional simulations.

[9]. This strategy is based on alternative error estimates and allows for a unified treatment of the unconditional convergence case and convergence with rates. In our future research we will try to extend the analysis to faster inner iterations than Landweber such as steepest descent or conjugate gradient methods.

6. Acknowledgment. Support by the German Science Foundation DFG under grant KA 1778/5-1 is gratefully acknowledged.

REFERENCES

- [1] A. B. Bakushinsky and M. Yu. Kokurin. *Iterative methods for approximate solution of inverse problems*. Springer, Dordrecht, 2004.
- [2] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Problems*, 20(5):1411–1421, 2004.
- [3] C. Clason and B. Jin: A semi-smooth Newton method for nonlinear parameter identification problems with impulsive noise, *SIAM J. Imaging Sci* 5:505–538, 2012.
- [4] M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numer. Math.* 72:21–37, 1995.
- [5] T. Hein and B. Hofmann. Approximate source conditions for nonlinear ill-posed problems – chances and limitations. *Inverse Problems*, 25:035003 (16pp), 2009.
- [6] T. Hein and K.S. Kazimierski. Accelerated Landweber iteration in Banach spaces. *Inverse Problems*, 26:055002 (17pp), 2010.
- [7] T. Hein and K.S. Kazimierski. Modified Landweber iteration in Banach spaces - convergence and convergence rates *Numerical Functional Analysis and Optimization* 31 (10), 2010, 1158 - 1189
- [8] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23(3):987–1010, 2007.
- [9] B. Kaltenbacher and I. Tomba. Convergence rates for an iteratively regularized Newton-Landweber iteration in Banach space. *Inverse Problems*, 29:025010 (18pp), 2013.
- [10] Q. Jin. Inexact Newton-Landweber iteration for solving nonlinear inverse problems in Banach spaces. *Inverse Problems* 28:065002 (14pp), 2012.
- [11] Q. Jin and L.Stals. Nonstationary iterated Tikhonov regularization for ill-posed problems in Banach spaces. *Inverse Problems*, to appear.
- [12] B. Kaltenbacher. Convergence rates for the iteratively regularized Landweber iteration in Banach space. *Proceedings of the 25th IFIP TC7 Conference on System Modeling and Optimization*, Springer, 2013, to appear.

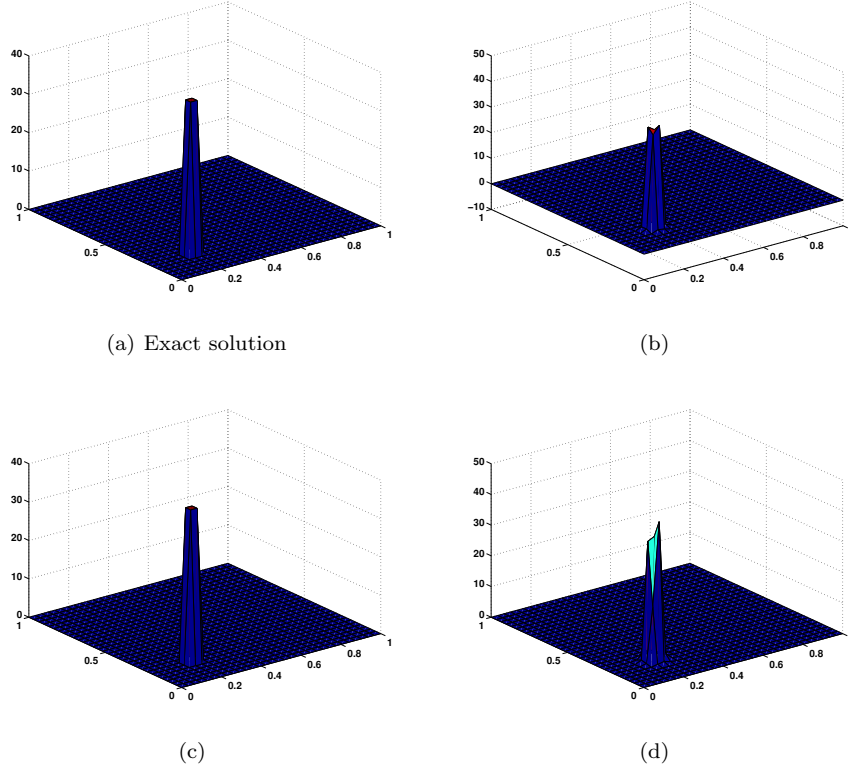


FIGURE 4. Numerical results for the 2-dimensional example: (a) is the exact solution; (b) is the solution corresponding to $\delta = 10^{-3}$ and $r = 2$; (c) is the solution corresponding to $\delta = 10^{-2}$ and $r = 2$; (d) is the solution corresponding to $\delta = 10^{-2}$ and $r = 10$.

- [13] B. Kaltenbacher and B. Hofmann. Convergence rates for the iteratively regularized gauss-newton method in Banach spaces. *Inverse Problems*, 26:035007 (21pp), 2010.
- [14] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Non-linear Ill-posed Problems*. de Gruyter, 2007.
- [15] B. Kaltenbacher, F. Schöpfer, and Th. Schuster. Convergence of some iterative methods for the regularization of nonlinear ill-posed problems in Banach spaces. *Inverse Problems*, 25, 2009. 065003 doi: 10.1088/0266-5611/25/6/065003.
- [16] A. Neubauer, T. Hein, B. Hofmann, S. Kindermann, and U. Tautenhahn. Improved and extended results for enhanced convergence rates of Tikhonov regularization in Banach spaces. *Appl. Anal.*, 89(11):1729–1743, 2010.
- [17] A. Rieder. On convergence rates of inexact Newton regularizations. *Numer. Math.* 88:347–365, 2001.
- [18] A. Rieder. Inexact Newton regularization using conjugate gradients as inner iteration. *SIAM J. Numer. Anal.* 43:604–622, 2005.
- [19] O. Scherzer. A modified Landweber iteration for solving parameter estimation problems. *Appl. Math. Optim.*, 38:45–68, 1998.
- [20] F. Schöpfer, A. K. Louis, and T. Schuster. Nonlinear iterative methods for linear ill-posed problems in Banach spaces. *Inverse Problems*, 22(1):311–329, 2006.
- [21] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. Kazimierski. *Regularization Methods in Banach Spaces*. de Gruyter, Berlin, New York, 2012.

- [22] Z.-B. Xu and G. F. Roach. Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces. *Journal of Mathematical Analysis and Applications*, **157** (1991), 189–210.

E-mail address: barbara.kaltenbacher@aau.at

E-mail address: ivan.tomba3@unibo.it